United States Department of Agriculture

Service Center Initiative (SCI)

Standard for Service Center Tabular Metadata

Prepared by the
Service Center Data Team

Abstract: Standard and minimum requirements for metadata (documentation of the data produced and used) to be provided for all Service Center development initiatives to populate the metadata repository.

Keywords: standard, metadata, tabular, repository

Published and distributed in ~~September 1~~ October 8 ~~July 16~~, 1999
(Revision of ~~Service Center Tabular Metadata Standard~~Standard for Service Center Tabular Metadata, September 1, ~~June~~ 1999)

Introduction

(This introduction is not part of SCI Std 008-~~01~~02, Standard for Service Center Tabular Metadata.)

A revision to SCI Std 008-01 was made. The working group list contained in SCI Std 008-01 was incorrect. SCI Std 008-02 has been updated with the correct working group list.

The Service Center Data Team developed this standard to guide Service Center development teams in understanding requirements for tabular metadata, and preparing that metadata for upload to the central repository. It specifies the circumstances under which metadata is required, and describes the specific requirements.

The individuals who contributed to the development of this guide are:

~~Glenn Bethel~~Dave Butler, (NRCS~~FSA~~) ~~Co-~~Leader
Tom Richards~~Carol Ernst~~, (FSA) Alt. ~~Co-~~Leader
Jim Schmidt~~Emil Horvath~~, (NRCS) Alt. ~~Co-~~Leader

John Bussone (FSA)	Mindy Gehrt (FSA)	Steve Springer (FSA)
Chuck Calcaterra (RD)	Lynette Harris (NRCS)	Lee Stieger (FSA)
Julie Degner (SOZA)	Jerry Kleiman (FSA)	Rich Verhoff (RD)
Dorn Egley (NRCS)~~Liz Cook (NRCS)~~		~~David Feld (RD) Lee Stieger (FSA)~~David
~~Randy English (NRCS)~~~~Feld (RD)~~		Craig Loddeke (RD) ~~Jill Schuler (NRCS)~~
	~~Kent Williams (FSA)~~Norma Westbrook (FSA)~~	

Figure 1—Working group list

Table of contents

This page is intentionally left blank.

STANDARD FOR SERVICE CENTER TABULAR METADATA

1. Overview

Metadata is simply information describing certain characteristics of a group or set of data. Development teams create this metadata to describe the requirements or characteristics of the data that will be created and/or used in the application systems being developed. Tabular data refers specifically to data that can be arranged in rows and columns, such as that which may be displayed in a spreadsheet or table.

The metadata pertaining to individual development projects provides a detailed view of the information associated with a specific application system. When this metadata is collected for all application systems and managed centrally through a repository, it provides a powerful tool for planning, managing and administering data on an enterprise level. In order to make the best use of this collection of metadata from across the organization, it is important that it be collected and structured in a manner that ensures that it is meaningful from an enterprise perspective. This requires ensuring that each project's metadata conforms to certain content and format standards.

1.1. Scope

This document describes the standards and minimum requirements for tabular metadata associated with Service Center development initiatives. This applies to new development projects, as well as redesign efforts.

NOTE—Most Geographic Information Systems (GIS) make provisions for the collection of a limited set of tabular data to describe each point, line or polygon feature drawn. This tabular data is often stored by the GIS system in a file available for access by users or by other information systems. For additional information on geospatial metadata refer to *Geospatial Feature Metadata Standard* [A1].

1.2. Purpose

This document defines the minimum requirements for metadata to be provided by Service Center teams working on development initiatives. Its purpose is to describe what is required, in terms of both metadata content and format, to a level of detail sufficient for development teams to clearly understand what information they are required to provide for upload to the central repository.

1.3. Acronyms and abbreviations

CASE	Computer-aided Software Engineering
FSA	Farm Service Agency
GIS	Geographic Information Systems
NRCS	Natural Resources Conservation Service

RD	Rural Development
SCI	Service Center Initiative
USDA	United States Department of Agriculture

2. Minimum metadata requirements

This standard defines the **minimum** requirement for metadata to be developed by each project. Some projects may collect additional information to document data requirements to a more detailed level. This section, however, describes the metadata that **must** be provided.

2.1. Data principles

The following principles apply to the definition and management of data and metadata, and were used to shape the minimum requirements and standards for tabular metadata.

1. Data is an asset with high monetary value, and must be managed in much the same way as other critical business area assets.
2. Business areas in which the data is created and used are ultimately responsible for the content, usage and distribution of any data.
3. Business areas must openly assign responsibility (stewardship) for every data asset it produces or relies upon, for meeting its mission. Every piece of data stored and used by an agency must have a steward responsible for its content and protection.
4. A data element can be shared or used by multiple functional business areas, groups, systems and subsystems. Changes to a data element can impact any or all of these categories of users.
5. The definition or true meaning of a data element is what makes the piece of data unique, not the name of the element.

2.2. Metadata categories

Minimum tabular metadata requirements are classified according to the following categories:

Business area	A portion of the United States Department of Agriculture (USDA) mission that relates to particular business activities performed within a Service Center. This may be a program area, a technical function, a managerial function, etc. In the organizational structure of an agency, this likely equates to a division, a major functional branch, a center or institute, a processing center, a scientific or business support center, etc. Metadata will be recorded for each business area that creates or uses electronic data.
System	An automated set of data and software that meets a specific business need. A system may consist of one or more

	software applications/subsystems. Metadata will be recorded for each system that stores or uses stored data.
Subsystem	A subset of a system, both data and software, that fulfills a particular requirement of the system. Metadata will be recorded for each subsystem.
Data element group	A set of concepts or things in the natural world whose properties and behavior all follow the same rules. Basically, a group equates to a file, a relational entity (table), an object class, a spatial feature type, and can even include aggregations of data into a document, a binary object (blob), a graphic, sound, or other multi-media data grouping, etc. Metadata will be captured for each data element group.
Data element	A discrete unit of data that is considered indivisible, has a specific definition, and is identified by a business name. Metadata is required for every data element.

2.3. Metadata requirements by category

Table 1 lists the required metadata elements to be recorded for each of the categories described in 2.2. Definitions for required metadata elements are provided in Appendix B.

Table 1—Required metadata

Category	Minimum metadata requirement
Business area	Business area name Business area definition Executive sponsor (organization, phone number~~#~~) Business area data steward (organization, phone number~~#~~)
System	System name System acronym System definition System manager (organization, phone number~~#~~)
Subsystem	Subsystem name Subsystem acronym Subsystem definition Subsystem version label Data element groups in the subsystem Business area name System name Data steward (organization, phone number~~#~~)

Category	Minimum metadata requirement
Data element group	Group name Group definition Data elements in the group Data steward (organization, phone number~~#~~)
Data element	Data element name Data element definition Data type Length Precision Permissible values Data steward (organization, phone number~~#~~)

NOTE—In Table 1, stewardship is identified by defining the responsible organization. Data stewardship responsibilities cover both the initial defining of data elements, and the subsequent physical data that populates the elements and element groups. The responsibilities of the Executive Sponsors and Data Stewards are described in additional detail in the *Service Center Data Administration Concept of Operations* [A2].

3. Metadata capture tools and process overview

~~Metadata capture tools and process overview~~

This section describes the tools recommended for the collection and an overview of the process for handing-off required metadata for upload to the central repository.

3.1. Metadata capture tools

Computer-aided Software Engineering (CASE) tools and other repository tools will provide the primary means for capturing metadata. Emphasis is being placed on tools that may reduce the workload associated with the collection and upload of required metadata.

In addition to those specified as minimum metadata requirements, there are additional metadata attributes that can be collected about data, some of which may be mandatory by certain CASE tools. Only the mandatory items specified in this document need to be uploaded to the central metadata repository. Additional attributes collected for specific development projects will be managed via the CASE tool, or other project documentation system.

For small projects, and for projects lacking CASE tools, other means will be made available for collecting required metadata.

3.2. Overview of metadata capture and hand-off process

The process associated with the capture of required metadata and the hand-off of that data for upload into the central repository is~~:~~

~~?~~ A data model depicting the information requirements associated with a development initiative is developed as part of the requirements definition.

~~?~~ The data model, along with associated metadata is entered into a CASE tool (or other electronic media, in certain circumstances.)

~~?~~ When the data model and metadata are complete (or stable) and have been validated by the stewards and business area experts, the electronic file(s) of the metadata are provided to the appropriate Service Center Data Manager.

~~?~~ A designated Service Center Data Manager reviews the model and metadata with regard to the system's integration into the overall enterprise, the potential impacts on data storage and transmission facilities and the completeness of the submission.

~~?~~ Once reviewed, the metadata is uploaded from the electronic file(s) provided to the central metadata repository in Kansas City.

~~?~~ Change control is established by the project over the submitted model and metadata. Essentially, the premise is that because users and other applications become dependent on data, all future changes to a data element or data element group will be announced and reviewed in advance. Typically, the Service Center Data Manager, and system managers and developers who are dependent on the data, participate in the review. Impacts of the proposed change to the enterprise are evaluated. Change control is described further in the document

Service Center Data Management Deliverables for Application Development Projects [A3].

Appendix A – Bibliography

[B~~A~~1] Geospatial Feature Metadata Standard.

[B~~A~~2] Service Center Data Administration Concept of Operations.

[B~~A~~3] Service Center Data Management Deliverables for Application Development Projects.

Appendix B – Definitions of required metadata elements

Table B.12 ~~–~~ ~~—~~Required metadata definitions

Element name			Definition	Domain
Business area			A portion of the USDA Mission that relates to particular business activities performed within a Service Center. Metadata will be recorded for each business area that creates or uses electronic data.	Business area name Executive sponsor Business area data steward
	Business area name		The label used to identify the business area within the organizational structure of the agency/agencies.	Free text
	Business area definition		Statement that describes the basic purpose and scope of the business area, and permits its differentiation from other business areas.	Free text
	Executive sponsor		The Executive Sponsor is a business-area manager who is accountable for the collection, management and use of data assets. The person may delegate day-to-day management responsibilities to a Data Steward, but ultimately has overall responsibility for the defining of data needed to support business-area operations, for the creation of software systems to collect and process the data, and all issues that deal with collecting, certifying and utilizing the data resources. In some cases this may be a shared responsibility between several business-area managers from different agencies.	Organization Organization phone number~~#~~
		Organization	The label used to identify the business area that is responsible for the metadata and data.	See footnote [1]

[1] The responsible organization is defined by specifying each level of the organization from agency down to the level that has the given responsibility.

Organization level	Is required?	FSA example	RD example	NRCS example
Agency abbreviation	Always required	FSA	RD	NRCS
Organization level 1 name	Always required	Financial Accounting	Rural Housing Service	Soil Survey and Resource Assessment
Organization level 2 name	As needed	General Ledger Accounting	Deputy Administrator Single Family Housing	Resource Inventory Division
Organization level 3 name	As needed		Single Family Housing Direct Loan Division	
Organization level 4 name	As needed			
Organization level 5 name	As needed			
Organization level 6 name	As needed			

Element name			Definition	Domain
		Organization Phone number~~#~~	A main phone number within the organization to be used to reach the Executive Sponsor.	Free text
	Business area data steward		A business-area expert designated by the Executive Sponsor to take day-to-day responsibility for all data collected and used to meet business-area needs. The data steward establishes business rules, defines data elements, identifies valid data values, establishes certification standards and establishes the completeness and availability of the data.	Organization Organization phone number
		Organization	The label used to identify the business area that is responsible for the metadata and data.	See footnote [1]
		Organization phone number~~#~~	A main phone number within the organization to be used to reach the data steward.	Free text
System			An automated set of data and software that meets a specific business need. A system may consist of one or more software applications/subsystems.	System name System definition System manager
	System name		The label used to identify the software system, i.e., the system name. System names must be unique.	Free text
	System acronym		The acronym of the system. System acronyms must be unique.	Free text
	System definition		Statement that describes the basic purpose and scope of the system, and permits its differentiation from other systems.	Free text
	System manager		The manager within the IT Organizational Unit responsible for the development of the system.	Organization Organization phone number~~#~~
		Organization	The label used to identify the organization that is responsible for development of the system.	See footnote [1]
		Organization phone number~~#~~	A main phone number within the organization to be used to reach the system manager.	Free text
Sub-system			A subset of a system, both data and software, that fulfills a particular requirement of the system.	Subsystem name Subsystem definition Business area name System name Data steward
	Subsystem name		The label used to identify the subsystem, i.e., the subsystem name. Subsystem names must be unique.	Free text
	Subsystem acronym		The acronym of the subsystem. Subsystem acronyms must be unique.	Free text
	Subsyste~~m~~m definition		Statement that describes the basic purpose and scope of the subsystem, and permits its differentiation from other subsystems.	Free text

Element name			Definition	Domain
	Subsystem version label		The version label of the subsystem that uniquely identifies a particular release of a subsystem's software or data.	Free text
	Business area name		The name of the Business Area sponsoring this subsystem.	Free text
	System name		The name of the System this subsystem belongs to.	Free text
	Data steward		A business-area expert designated by the business-area data steward or by the executive sponsor to take day-to-day responsibility for the data created and used by this particular subsystem. The subsystem data steward develops business rules, defines data elements, identifies valid data values and completes the metadata for the particular software subsystem.	Organization Organization phone number~~#~~
		Organization	The label used to identify the business-area that is responsible for the metadata and data.	See footnote [1]
		Phone ~~#~~number	A main phone number within the organization to be used to reach the data steward.	Free text
Data element group			A set of concepts or things in the natural world whose properties and behavior all follow the same rules. In the relational world this is known as an entity or entity type. In the object world this is known as an object class or object type. In the geospatial world this is known as a feature type or entity type. Data can also be grouped into a file, a document, folder, binary object, etc.	Group name Group definition Data elements in group Data steward
	Group name		The label used to identify the grouping. Data element group names are unique within a system. There are cases where distinct data element groups may have the same business name.	Free text
	Group definition		Statement that describes the basic content of the data element grouping, and permits its differentiation from other groupings.	Free text
	Data elements in the group		A listing of data elements contained in the group. Through the CASE tool, this may be implemented as a set of internal linkages within the tool.	Free text or constructed linkages within a tool which allow for the viewing of group members upon request.

Element name			Definition	Domain
	Data steward		A business-area expert designated by the business-area data steward or by the executive sponsor to take day-to-day responsibility for the data in this particular data element grouping. The data steward develops business rules, defines data elements, identifies valid data values and completes the metadata for the particular data element grouping.	Organization Organization phone #number (See data steward under "Subsystems" for definitions of these elements)
Data element			A discrete unit of data that is considered indivisible, has a specific definition and is identified by a business name.	Data element name Data element definition Datatype, Length, Precision permissible values Data steward
	Data element name		The label used to identify the element. Data element names are unique within a system. There are cases where distinct data elements may have the same business name.	Free text Names are defined in accordance with the USDA Service Center Data Naming Standard for "business" names.
	Data element definition		Statement that expresses the essential meaning of the item and permits its differentiation from other items.	Free text
	Datatype		Datatype is the format used to store the data.	[2] Examples: Character, numeric, integer, packed decimal, floating-point, text, sound, video, binary, etc.
	Length, precision		Length: the total number of digits (fractional precision included) for fixed decimal datatypes. The maximum length in characters (or bytes) for other datatypes. Precision: the number of digits to the right of the decimal for fixed decimal datatypes.	See footnote 2.
	Permissible values		A description of the permissible values of the domain. The set of values can be specified as a range, as an enumerated list or as a reference to the source which defines the permissible values.	[3] A range, an enumeration list or a reference.

[2] When metadata is defined in a CASE tool, the datatype, length and precision should be specified using the *logical* model features of the CASE tool. When metadata is defined outside of a CASE tool, datatype, length and precision are specified as follows:

Character length
Decimal length, Precision
Integer
Float
Sound, Video
Text, Binary
Date, Time, Timestamp

[3] Permissible Values Format

Range: Lower and upper limit separated by a "-"

Element name			Definition	Domain
	Data steward		A business-area expert designated by the Business-area Data Steward or by the Executive Sponsor to take day-to-day responsibility for this data element. The data steward develops business rules for populating this element, identifies valid data values, and completes the metadata for the particular data element.	Organization Organization phone #number (See data steward under “Subsystems” for definitions of these elements)

Enumeration: "value – description" pairs separated by a <tab>; one per line
Reference: Reference document title or other free form text

Appendix C – Administrative metadata elements

The following table provides an overview of administrative metadata elements. This information is identified, recorded and maintained as part of the data administration process.

Table ~~3~~C.1 —— Administrative metadata definitions

Name	Definition	Domain	Comments
Data element standardization level	A designation of the position in the standardization life cycle of a registered item. This is essentially a quality indicator of the metadata being maintained. Once under change control this is considered a final designation of the quality of the metadata.	[4]Incomplete, Legacy, Recorded, Certified, Standard, Retired	Only applies to data elements.
Administrative status	A designation of the position in the processing life cycle of a registered item.	[5]Work in process, Under change control, Operational, Inactive	"Work in process," "Under change control," and "Inactive" apply to data elements, data element groups and subsystems. "Operational" applies only to subsystems. Systems and business areas, which are essentially static definitions, do not have an administrative status.
Identifier	The label that uniquely identifies the registered item, e.g., data element, data element group, etc.		Assigned as part of the data administration process.
Last modification date	The timestamp of most recent posting of the metadata to the repository.		Assigned as part of the data administration process.
Modified by	The label that identifies who was responsible for the most recent posting to the repository.		Assigned as part of the data administration process.

[4] Data Element Standardization Level

Incomplete	Does not contain all mandatory metadata. This level is not allowed for elements that are under change control.
Legacy	Metadata has been loaded for a legacy system as is. Some metadata may be incomplete. The metadata may not pass quality criteria. There is no intent to "fix" the metadata.
Recorded	All mandatory metadata has been recorded but the contents may not meet quality criteria. This is a final designation; i.e., there is no intent to do a quality review.
Certified:	All mandatory metadata has been recorded and the contents meet quality criteria. This is a final designation. However, the data element may become a candidate to become a standard element.
Standard:	A certified item that has been designated as the preferred item. It may be unique within the repository or it may be the preferred item among similar items.
Retired:	The data element is no longer used.

[5] Registered Item Administrative Status

Work In Process:	metadata may still be changed before going under change control
Under Change Control:	metadata is now under change control
Operational:	the subsystem is now operational
Inactive:	the subsystem is now inactive, i.e. it is no longer in use; may also apply to data elements and data element groups that are no longer used by any active subsystem